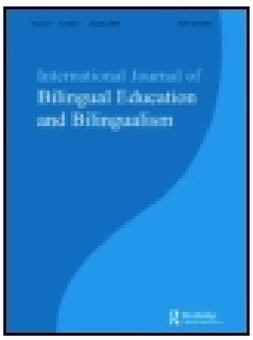
Linguistic consequences of toing and froing



International Journal of Bilingual Education and Bilingualism

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rbeb20

Linguistic consequences of toing and froing: factors that modulate narrative development in bilingual returnee children

Maki Kubota, Vicky Chondrogianni, Adam Scott Clark & Jason Rothman

To cite this article: Maki Kubota, Vicky Chondrogianni, Adam Scott Clark & Jason Rothman (2021): Linguistic consequences of toing and froing: factors that modulate narrative development in bilingual returnee children, International Journal of Bilingual Education and Bilingualism, DOI: 10.1080/13670050.2021.1910621

To link to this article: https://doi.org/10.1080/13670050.2021.1910621

INTERNATIONAL JOURNAL OF BILINGUAL EDUCATION AND BILINGUALISM
https://doi.org/10.1080/13670050.2021.1910621

Linguistic consequences of toing and froing: factors that modulate narrative development in bilingual returnee children

Maki Kubota [a], Vicky Chondrogianni[b], Adam Scott Clark[c] and Jason Rothman[a,d]

[a]Department of Language and Culture, UiT The Arctic University of Norway, Tromsø, Norway; [b]School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh; [c]Independent Scholar, Toronto, Canada; [d]Department of Language and Education, Universidad Nebrija, Tromsø, Norway

ABSTRACT

This longitudinal study examined the development of narrative micro- and macrostructure in Japanese-English bilingual returnee children. Returnees are children of immigrant families who move to a foreign country, spending a significant portion of their formative developmental years in the foreign majority language context before returning to their native language environment. The returnees did a narrative task in both their L1 (Japanese) and L2 (English) immediately upon their return to their native language environment and a year after. The results showed no aggregate significant changes in L1 or L2 micro- and macrostructure over time. However, at the individual level, the degree of maintenance of L2 microstructure was modulated by L2 exposure. That is, children who continued to receive L2 exposure better maintained their English microstructure (i.e. Type-Token Ratio and Verbs per Utterance) despite being re-immersed in the L1 environment. In terms of their Japanese, the age of return to the L1 environment and relative proficiency predicted the development of their Japanese microstructure (i.e. MLU, Fluency, Type-Token Ratio) and macrostructure. Our study is the first to track both languages of bilingual returnee children over time, revealing that different background variables affect the change in returnee children's L1 and L2 narrative abilities.

ARTICLE HISTORY
Received 7 October 2020
Accepted 15 March 2021

KEYWORDS
Returnee children; narrative; microstructure; macrostructure; longitudinal

1. Introduction

Returnees are a special group of bilinguals who, in the typical case, move from their native birth country and spend a significant portion of their formative years in an immigrant setting before returning 'home'. One can also be a returnee if born to a migrant family within an immigrant country. Under the latter scenario, returnees can be simultaneously (2L1) bilinguals of both the home (or heritage language) as well as the societal majority language, depending on the age/timing of significant exposure to the latter. For returnees, exposure to their first language (L1) is likely to be, on a spectrum, significantly reduced, potentially to only within the family unit or a small diaspora minority community. Within the wider community, they become exposed to (and likely dominant in) the majority environmental language (Flores and Snape forthcoming). Crucially what makes returnees distinct from a typical heritage language bilingual – children of immigrants who speak a minority language at home that is distinct from the societal majority language (see

CONTACT Maki Kubota makikubota5@gmail.com
Supplemental data for this article can be accessed online at https://doi.org/10.1080/13670050.2021.1910621.

e.g. Montrul 2010; Polinsky 2018; Rothman 2009) – is their moving to ('returning' to) their ancestral/heritage country of origin at some point in later childhood or early adulthood.

Typical heritage language bilinguals tend to develop and maintain skewed dominance in favor of the societal majority language over time. While living abroad, returnees display all the signs of typical heritage bilingualism, however, they are a special subset of heritage bilinguals because they have the opportunity to shift dominance (back) towards their cultural language Therefore, they are often at risk of continuing to develop and/or maintain the language of the host country in which they were likely dominant prior to returning (Flores and Snape forthcoming). In the case of returnees, the status of the majority and the minority language reverses – what was formerly the minority or heritage language in an immigration setting becomes (again) the majority language of the community, while the former majority language becomes a minority language (Flores 2019). This environmental change can lead to two major linguistic consequences, namely, *second language (L2) attrition* and *heritage language reversal*. *Second language attrition* is a process in which returnees' L2 knowledge (competence in the majority language of the host country) becomes less accessible or is modified to some extent as a result of re-immersion in their native language environment and reduced contact to the L2. *Heritage language reversal*, on the other hand, concerns the developmental changes that occur in one's native/heritage language once it reverses from being the minority language to the dominant environmental language (Flores and Snape forthcoming).

In addition to the fact that research on returnee children is scarce, the majority of existing studies either focus on the process of *second language attrition* or *heritage language reversal*. To date, no study prior to the present one has examined *both* processes simultaneously among the returnee population. Revealing the linguistic consequences of returnees is especially important in the current context of global mobility where language development in the native and additional languages for millions of children is dynamic and fluid. For example, in Europe alone, in excess of one million EU citizens return 'home' to their (distinct EU) birth country or ancestral homeland annually (Eurostat 2019). Understanding the consequences of the returnee context for development and maintenance in both languages is a necessary first step in assessing and, ultimately, being able to meet the educational/interventional needs of returnee children.

It would be misleading to conceive of all returnees as having a unifying set of circumstances to their background simply because they qualify as returnees. Just like any set of language learners – monolingual, bilingual or multilingual – significant differences at the individual and group levels exist. Let us take the case of Japanese returnees as an example. Japanese returnees could be labeled as either 'kikokushijo' i.e. children of Japanese expats who migrate overseas for work assignments and return to Japan, or 'nikkei' i.e. Japanese-descent immigrants born in a different country (particularly from Brazil or Peru) who often move to Japan to seek opportunities for employment in early adulthood or as older children with their parents. While the former group of returnees is generally younger, educated, and socioeconomically privileged, returnees from the latter group are often older, economically disadvantaged, and have low-skilled jobs (Kubota, 2012). What is common among them, however, is the fact that they receive little language support when they re-integrate into the Japanese environment (Kubota, 2012). What they are also likely to share in common is the need for intervention. In a large-sclae survey study in Japan, 87.5% of teachers of Nikkei Brazilian children have reported them to have various language-related problems in class (Sekiguchi 2003). In similar vein, Yoshida et al. (2009) tested 512 'kikokushijo' returnees and found that many of them experienced language-related issues upon returning to Japan, both in terms of maintaining their L2 and catching up to their peers in their L1. Thus, despite differences that are sure to pertain to various cohorts of returnees and individuals, primary research of the type we present here is likely to offer some generalizable insights.

The current study aims to track the development of both the L1 (Japanese) and L2 (English) of Japanese returnee children over the course of the first year of their return to Japan. We measured their narrative productive abilities, which assess multiple linguistic features including macrostructure (the global organization of a story) and microstructure (usage of complex syntactic structures and

specific types of words). Narratives are a good point of departure to examine in the returnee population as they correlate with other educational outcomes such as writing (Spencer and Petersen 2018), reading (Bailey et al. 2018), mathematical abilities (O'neill, Pearce, and Pick 2004), and critical thinking and learning motivations (Aktas and Yurt 2017; Yang and Wu 2012). Most importantly, we investigated whether (and which dimensions of) bilingual experience (e.g. exposure, age, length of residence, proficiency, age of onset) predict the changes in returnee's L1 and L2 narrative abilities over time.

1.1. Narrative abilities

Narratives are commonly used for studying child language development in monolinguals and bilinguals, as they provide a wide range of information about children's linguistic knowledge and are found to be good predictors of literacy and academic outcomes (Miller et al. 2006; To et al. 2010). Two measures are usually elicited through narrative production tasks: microstructure and macrostructure. Macrostructure is concerned with the overarching structure of the story and its organizational schema, while microstructure is related to linguistic knowledge at the sentence and word level such as Mean Length of Utterance (MLU), lexical diversity (e.g. type token ratio or number of different words), and syntactic complexity (e.g. verbs per utterance or degree of sentence subordination) (Justice et al. 2006).

Within the bilingual literature, it has been suggested that macrostructure is universal across languages while microsturcture is language specific (Méndez et al. 2018). The idea here is that underlying knowledge of constructing a coherent story grammar may rely more on cognitive abilities that are shared between languages (Gagarina et al. 2016; Iluz-Cohen and Walters 2012; Lucero 2015; Squires et al. 2014), whereas knowledge of vocabulary and grammar are highly dependent and specific to each language and, therefore, may not easily transfer across languages (Iluz-Cohen and Walters 2012; Rodina 2017; Simon-Cereijido and Gutiérrez-Clellen 2009).

Not all studies have yielded consistent findings on the relationship between macro- and microstructure within and across languages. For instance, Westerveld (2014) found no correlations of macrostructure performance across languages in a narrative retell task. Lucero (2015) showed that some variables of both micro- and macrostructure were associated across languages. Such discrepancies across studies may be ascribed to individual differences in bilingual experience – performance of micro- and macrostructure may be modulated by combination of various factors such as age, language proficiency, and input/exposure. Rodina (2017), for example, argues that language exposure to each of their language is important for bilingual children's development of microstructure. In a similar vein, Govindarajan and Paradis (2019) found that length and quality of L2 exposure predicted micro- and macrostructure elements in five to six year-old L2 English bilingual children. Moreover, measures of vocabulary which are often used as indices of language proficiency, predicted how well one can tell a coherent story in the speakers' heritage language (i.e. Spanish), but not in their majority language (i.e. English) (Uccelli and Páez 2007). Although vast amount of studies on narratives exist in the bilingual literature, factors that modulate these abilities still remain underxplored and are open to further investigation (Pesco and Bird 2016).

1.2. Second language attrition

Narratives are in fact one of the most widely used tools to elicit language production in the studies of second language attrition among returnee children (Hansen-Strain 1990; Reetz-Kurashige 1999; Taura and Taura 2000; Tomiyama 1999, 2000, 2008; Yoshitomi 1999). For instance, the following studies, among other tasks, all used the same word-less picture book(s) (i.e. Mayer's frog series: Mayer 1969) to examine the degree of L2 English attrition in Japanese returnee children. Most of these studies, however, focused primarily on L2 microstructure, and no study to date has looked at the attrition of macrostructure in returnee children.

Reetz-Kurashige (1999) conducted a cross-sectional study on 18 returnees (age range from 6;5 to 13;7), who had been back to Japan for 3–22 months. The results showed no difference between the control and the returnee group in terms of lexical diversity (i.e. type-token counts). However, there were some shifts observed in their use of tense and aspect and common morphological errors were found such as dropping the *-ed* morpheme in the past tense and the *-s* for third person singular. Tomiyama's (1999, 2000) study followed the English attrition of a Japanese returnee child (8,0 at the onset of the study) longitudinally for 33 months. The data were collected once a month through a number of instruments, including free conversation, the Peabody Picture Vocabulary Test (PPVT), the Bilingual Syntax Measure (BSM), and a narrative task. Fluency measured by the amount of code switching, long pauses, and other compensatory strategies were observed in the first 19 months of the study. However, there were no signs of attrition in the child's receptive lexicon measured by PPVT. In the 20–33-month period after returning to Japan, the child showed some signs of attrition in morphology (plural and past irregular morphemes) and syntax (modification of nouns, preposition, and relativization).

Finally, a study by Taura (2001) explored the attrition of English in 21 Japanese high school returnees who were educated in an English-speaking environment for more than three years. The total number of words produced, type-token ratio, dysfluency rate (i.e. total pause duration divided by the total time spent on speech production), and syntactic complexity rate (i.e. the number of embedded clauses divided by the number of T-units) were analyzed. The results showed that children who appear to have surpassed a certain level of 'threshold' in L2 proficiency produced the most complex sentences, with the highest fluency rate and largest vocabulary size. Moreover, similar to Tomiyama's (1999, 2000) findings, Taura's results also demonstrated a fluctuation in the participants' L2 English knowledge after re-integrating into the Japanese environment.

Taken together, all of these findings show that regardless of the limited exposure that Japanese returnee children receive in their L2, they generally tend to maintain their L2 knowledge but exhibit some signs of attrition in various linguistic aspects and properties. The factors that predict the (limited) changes we see in the Japanese returnee children's L2 still remain uncertain, given that most studies on L2 attrition are qualitative, involving a small number of participants, and are limited in number.

1.3. Heritage language reversal

As alluded to above, the linguistic consequences of returning are not limited to the L2, but also affect the native, yet formerly heritage language, as it, once again, becomes the societal, majority language. These results, among other things, in significantly increased L1 input, differential types of (formal) exposure, variety of interlocutors and opportunities for usage. Indeed, a limited number of studies have examined the changes that occur in returnees' L2 upon return to the homeland. However, research which specifically focus on the 'other' L1/HL is even more scarce, with only a handful of empirical studies investigating this process. Treffers-Daller et al. (2016) looked at the use of lexical collocations among Turkish returnees from Germany, Turkish HSs in Germany, and Turkish monolinguals. Their findings revealed that even returnees who had been back for only a year in Turkey demonstrated better performance in the choice of noun-verb collocations than the HSs. However, only the returnees who had been back to Turkey for seven years showed comparative performance to that of their monolingual peers. Once the returnees spend a significant amount of time in monolingual environments of the heritage language, they appear to catch up to the monolinguals on various aspects of the language such as syntactic embeddings (Treffers-Daller, Özsoy, and Van Hout 2007), motion event cognition (Daller, Treffers-Daller, and Furman 2011), and general proficiency (Daller and Yıldız 1995). The only study to our knowledge that has explored the variables that affect heritage language reversal is by Flores and Rato (2016), which investigated global accents in Portuguese returnees from Germany. They found that the age at which the HSs/returnees emigrated to Germany (but not length of re-immersion in Portugal) predicted the variability in their

ratings of Portuguese global accents. Despite the fact that tracking returnees' heritage language during childhood can provide us with valuable information regarding the impact of input quality and quantity on native language development, no study to date has examined this process in a group of child returnees in a longitudinal manner (see Flores 2020 for further discussion).

1.4. Present study

The current study has the following two main research questions:

(1) To what extent do linguistic changes occur in returnee's L1 and L2, starting from the point of return to their native language environment?
(2) What factors may explain potential changes/development that take place during the initial stages of re-integration into the homeland?

As we explain in detail below, the present study examines Japanese returnee children, all of whom were born in Japan and, therefore, are second language learners of the majority language in which they lived, English, prior to returning to Japan. Under the scenario of our participants, we predicted that their L1 Japanese microstructure would increase and their L2 English microstructure decrease over time (or there will be a trend towards such patterns). Alternatively, we expected a significant increase in both L1 and L2 macrostructure, since this is argued to be a skill that is universal across languages and develops generally with age. Based on previous literature, we focused on five bilingual factors that are good candidates for potential modulators of individual level changes in L1 and L2: age of return, exposure, length of residence in the L2 environment, age of L2 onset, and relative proficiency at the onset of return. We included all five factors in the statistical models to explore their predictive power, not least due to the dearth of relevant literature from which reasonable expectations could have been made *a priori*. We were not able to directly compare the narrative abilities of the L1 to the L2, since we used different wordless picture book for each language which involve different number of characters, events, and pictures. However, comparing narrative abilities between the two languages of a bilingual has been extensively examined in the field (Altman et al. 2016; Gagarina 2012; Govindarajan and Paradis 2019; Hao et al. 2019; Iluz-Cohen and Walters 2012; Kapalková et al. 2016; Kunnari, Välimaa, and Laukkanen-Nevala 2016; Lucero 2015; Méndez et al. 2018; Roch, Florit, and Levorato 2016; Rodina 2017; Squires et al. 2014; Uccelli and Páez 2007) and thus is not the main purpose of our study. Conversely, we were interested in investigating which factors modulate the changes in L1 and L2 narrative abilities in a rare(ly studied) circumstance in which the dominance of the language reverses due to shifts in the environment.

2. Methodology

2.1. Participants

This study (project title: L2 attrition and L1 acquisition in Japanese-English bilingual children) was approved by The University of Edinburgh PPLS Ethics Committee. Thirty-eight Japanese-English bilingual returnee children initially participated in the study. Four participants' data were lost due to technical issues. An additional two children dropped out in the second round of testing and were thus removed from the final sample, resulting in 32 participants in total. The bilingual children (19 female and 13 male) were all born in Japan, have both parents as native speakers of Japanese and acquired English upon arrival to an English-speaking environment in early childhood. All the bilingual children had very limited exposure to English, if any, before leaving Japan. The only prior English exposure they would have had would have been through weekly language classes at a private institution or at their elementary school. The participant information of the bilingual

Table 1. Summary of bilingual participant information; Incubation period indicates the time elapsed between children's return to Japan and the first test session

	Mean	SD	Min	Max
Age at Time 1 (in months)	116.4	16.72	91.2	156.0
Age at Time 2 (in months)	129.6	17.04	8.6	14.0
Age of L2 immersion onset (in months)	61.2	31.2	14.4	116.4
Length of L2 residence (in years)	4.1	2.0	2.0	9.7
Incubation period (in years)	0.33	0.16	0.17	0.81

returnees is provided in Table 1. Incubation period indicates the time elapsed between children's return to Japan and the first test session.

These bilingual children attended schools with English as a medium of instruction in a foreign country. Half of the participants had lived in a country where English is the majority language (e.g. USA, UK, Australia), whereas the other half had lived in a country where English is not the official language (e.g. Malaysia, France, Netherlands). We made sure to control for this environmental difference by including Country (English is the official language of the country vs. English is not the official language of the country) as a control variable in all of our statistical models.[1] The children who were living in a country where English was not the official language were exposed to the dominant language of the country, but the parents all reported that their children could not hold a conversation using that language. Participants were recruited through Japan Overseas Educational Services, an organization that offers support for returnee families and children. All children were enrolled in a Japanese school upon their return to Japan and were educated under the curriculum set by the Japanese Ministry of Education.

2.2. Instruments

2.2.1. Language background questionnaire

In order to quantify language exposure for each language, the Bilingual Language Experience Calculator (BiLEC) (Unsworth 2016) was administered to the parents individually by the first author. Table 2 illustrates the quantified language exposure when they lived abroad (Abroad) and when they returned to Japan (Japan), as well as the difference in exposure between foreign and Japanese environment.

As shown in Table 2, it appears to be the case that the children were receiving balanced exposure in their L1 Japanese and L2 English when they lived abroad (55.3% and 44.7% respectively). On average, their English exposure decreased more than 40% since their return to Japan and consequently the inverse (increase) is true of their Japanese exposure. This difference in English exposure between the foreign environment and Japan is used later as a predictor (labeled as 'Exp_diff') on development in microstructure and macrostructure.

2.2.3. Proficiency task

A category verbal fluency task was used to measure children's relative language proficiency across both languages. Two semantic categories were tested: (a) fruits and vegetables and (b) animals. The

Table 2. Summary of BiLEC variables split by language and time; 'Abroad' indicates language exposures of when the children lived in a L2 majority language environment and 'Japan' indicates exposures of when the children returned to Japan; 'Difference' indicates the difference in exposure to each language between the foreign and Japanese envrionment (the numbers are in percentages).

	English (L2)			Japanese (L1)		
	Abroad	Japan	Difference	Abroad	Japan	Difference
Mean	44.7	4.2	−40.5	55.3	95.8	+40.5
SD	11.9	5.5	10.0	11.0	5.5	6.5
Min	26.5	0	−26.5	17.5	28.0	+15.0
Max	82.4	20.5	−58.0	61.0	92.4	+60.0

participants were asked to name as many things that belong to each category in one minute. Half of the bilinguals named animals in English and fruits and vegetables in Japanese, and vice-versa for the other half. The order of the language was also counterbalanced. The instructions were given in English for the English task and in Japanese for the Japanese task. Their responses were audio-recorded, and the total number of correct words was calculated. Repeated words were omitted from the analysis. The difference in the total number of unique words between English and Japanese at first round of testing (i.e. at the onset of return to Japan) were used for further analyses as a measure of their relative proficiency. Higher values indicate stronger proficiency in Japanese.

2.2.2. Story-telling task

We used two wordless picture books to elicit children's narratives: *Frog on his own* (Mayer 1973) and *Frog, where are you?* (Mayer 1969). *Frog on his own* was used to elicit Japanese narratives, whereas *Frog, where are you?* was used to elicit English narratives. *Frog on his own* consists of 30-pages of illustrations with 12 different characters and *Frog, where are you?* includes 29 illustrations with 7 different characters.

2.3. Procedure

The story-telling task was administered in a quiet room and the participants were seen individually by a Japanese-English bilingual researcher. The test session was conducted at the participant's home or in Japan Overseas Educational Services (JOES) classrooms. Half of the participants were asked to tell the story in Japanese first and then in English and vice-versa for the remaining half. The instructions were given in English for the English task and in Japanese for the Japanese task. The participants were allowed to first read through the book and prepare their story. They were told that they were being recorded so that other children could listen to their explanation of the story. Their responses were recorded on a voice-recorder and later transcribed by five research assistants. The story-telling task was administered twice: in the summer of 2016 and in the summer of 2017. Identical measures were taken for the second test session, including the order of the language task and the language of instruction.

2.4. Coding

2.4.1 Microstructure

We transcribed Japanese and English narratives according to the CHAT system (MacWhinney 2000) and analyzed them using Computerized Language Analysis (CLAN). Both Japanese and English transcriptions were first segmented into communication units (C-units), defined as an independent clause plus its modifier (Loban 1976). Three native speakers of English transcribed the English data and two Japanese-English bilinguals transcribed the Japanese data. The reliability among the three native English speakers who transcribed the English data was 93%, while the reliability between the two Japanese-English bilinguals who transcribed the Japanese data was 90%. Disagreements were resolved by the main researcher. Four different microstructure elements were derived: Mean Length of Utterance in word tokens (MLU), Type-Token Ratio (TTR), Typical Disfluency (TD), and Verbs per Utterance (VU). TD is the sum of phrase repetitions, word revisions, phrase revisions, pause counts, phonological fragments, and filled pauses over total words (MacWhinney 2000). VU roughly corresponds to clauses per utterance, and indexes grammatical complexity (MacWhinney 2000).

2.4.2 Macrostructure

We coded the macrostructure elements by adapting the rubrics established by Squires et al. (2014) for the *Frog on his own* story and by Hao et al. (2019) for the *Frog, where are you?* Story (see Supplementary Material). Both rubrics are based on the macrostructure subscale of the Monitoring Indicators of Scholarly Language (MISL) (Gillam et al. 2017) and contains seven elements: character,

setting, initiating event, internal response, plan, action, and consequence. Each element scored on a scale from 0 to 3, yielding a total macrostructure score of 21. Two Japanese-English bilinguals coded the macrostructure for both languages. Twenty percent of the narrative samples were randomly selected for independent coding. The inter-rater reliability for the English macrostructure coding was 98% and the reliability for Japanese macrostructure coding was 94%.

2.5. Data analysis

We ran mixed effect models to examine the change in micro- and macrostructure over time (from Round 1 to Round 2) and to see what dimensions of bilingual experience predict this change. For all models, Country (English is the official language of the country vs. English is not the official language of the country) was included as a means of controlling for variances that arise due to differences in the language environment. For the English data, we included age at return/age at Round 1 (Age), English exposure since they came back to Japan (L2_exposure), difference in L2 exposure between foreign and Japanese environment (Exp_diff), and length of residence (LoR) as predictors and Subject as random intercept. As for the Japanese data, we included age at Round 1 (Age), difference in L1 exposure between foreign and Japanese environment (Exp_diff), length of residence (LoR), and relative proficiency as predictors and Subject as random intercept. We included L2 exposure in the English model (but not in the Japanese model) and Relative Proficiency in the Japanese model (but not in the English model), to avoid issues of collinearity since these two variables highly correlated ($r = -.57$, $p > .001$). L2 exposure was favored over Relative Proficiency in the English model since past literature has shown L2 exposure to affect the rate of English attrition in returnee children (Kubota, Chevalier, and Sorace 2020a). However, L2 English exposure is not expected to affect the change in Japanese micro- and macrostructure over time, and thus, we included Relative Proficiency (in place of L2 exposure) in the Japanese models. The same is true of Age of L2 onset (AoA). Since AoA correlated highly with LoR ($r = -.85$, $p > .001$) and also moderately with Age ($r = .42$, $p = .01$), we excluded it from all models. This decision was also based on previous work on returnees which found LoR and Age of return to the L1 environment to be key factors in predicting the rate of L2 attrition (Flores 2010, 2020). Models were fit in R (R Core Team, 2013) with the package 'lme4' (Bates et al. 2015). We performed mixed model ANOVA tables via likelihood ratio test using the afex package (Singmann et al. 2015) on the output of lmer models. The afex package automatically applies sum coding to categorical variables. The reference level was set to 'Round 2' for the Round variable. All the continuous predictors were centered around the mean.

3. Results

We will first present the results of the verbal fluency task used to measure relative proficiency, followed by the results of microstructure and macrostructure in English, and finally the results of microstructure and macrostructure in Japanese.

3.1 Relative Proficiency

The children performed a verbaly fluency task and we computed the difference in scores between Japanese and English performance as a proxy for relative proficiency. The results of the verbal fluency task is presented in Table 3. A linear mixed effects model with Language (Japanese, English), Round (Round 1, Round 2), and Country (English as official language vs. non-official language) as fixed effects and Subject as random effects revealed that there were no significant effect of Round ($E = 1.54$, $t = 1.80$, $p = .07$), Language ($E = 1.59$, $t = 1.87$, $p = .06$) or interactions between Language and Round ($E = .79$, $t = .65$, $p = .51$). This indicates that children performed similarily in English and Japanese, and the scores in both languages did not increase from Round 1 to Round 2. However, there was a main effect of Country ($E = 3.33$, $t = 2.90$ $p = .006$), indicating that children who lived abroad where

Table 3. Summary of verbal fluency performance split by Language and Round; 'Round 1' indicates children's performance at first round of testing and 'Round 2' indicates their performance at second round of testing. Relative proficiency is the difference between Japanese and English performance. The numbers indicate how many unique items (animal or vegetable and fruits) the children were able to name within one minute.

	L2 English		L1 Japanese		Relative Proficiency	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
Mean	13.59	15.22	15.18	17.53	1.59	2.31
SD	4.42	3.29	5.13	5.41	5.05	5.39
Min	8	7	8	10	8	15
Max	27	22	34	31	−12	−8

English was not the official language performed better than children who lived in a country with English as the official language. This effect is however most likely cofounded with age, since children from the former group were older (mean = 124.15 months) than the latter group (mean = 109.16 months). Although dominance in Japanese appears to increase over time, there was no statistical change in relative proficiency from Round 1 to Round 2 ($t(60) = -.67$, $p = .50$).

3.2. English

3.2.1. Microstructure

The descriptive statistics of the English microstructure are presented in Table 4.

Overall, there appears to be no change in all microstructure elements from Round 1 to Round 2. This is confirmed by the results from the linear mixed effects model; no main effect of Round was found for Mean Length of Utterance (MLU) ($E = -.06$, $t = -.77$, $p = .44$), Type-Token Ratio (TTR) ($E = -.003$, $t = -.78$, $p = .44$), Typical Disfluency (TD) ($E = -.15$, $t = -.34$, $p = .73$), and Verbs per Utterance (VU) ($E = -.002$, $t = -.14$, $p = .88$). As for the effects of background variables on the change in microstructure over time, no significant interactions were found between Round and background variables for MLU (p's > .18) and TD (p's > .23), but a significant interaction was found between Round and Exp_diff for TTR ($E = .01$, $t = 2.22$, $p = .03$) as well as VU ($E = .04$, $t = 2.46$, $p = .02$). There was also no main effect of Country for all models (p's > .15).

The difference in L2 exposure between foreign and Japanese environment (Exp_diff) modulated the changes in TTR and VU from first to second round of testing. Figure 1 Panel (A) illustrates the interaction between Round and Exp_diff for TTR. Visual inspection shows that returnee children who experienced more reduction in L2 exposure started out with higher TTR immediately upon their return to Japan, and then experienced the greatest drop in TTR after a year of immersion in the Japanese environment. Simple regression analysis of Exp_diff on TTR at each time point revealed a significant effect of Exp_diff on TTR at the second round ($E = -.01$, $t = -2.1$ $p = .04$) but not at the first round of testing ($E = .01$, $t = 1.23$, $p = .22$).

Table 4. Summary of the English microstructure elements split by Round; 'Round 1' indicates children's performance at first round of testing and 'Round 2' indicates their performance at second round of testing.

	Mean	SD	Min	Max
Mean Length of Utterance				
Round 1	7.29	1.14	4.75	9.80
Round 2	7.47	.87	5.81	9.48
Type-Token Ratio				
Round 1	.32	.04	.20	.43
Round 2	.33	.03	.25	.45
Typical Disfluency				
Round 1	9.19	6.58	.67	33.11
Round 2	9.42	6.80	0	30.65
Verbs per Utterance				
Round 1	1.19	.17	.84	1.61
Round 2	1.21	.15	.80	1.66

Figure 1 Panel (B) shows the interaction between Round and Exp_diff on VU. Here, we see a similar pattern to the results obtained in TTR – continued exposure to the L2 contributed to an increase in VU, while experiencing more than around 45% of reduction in L2 exposure affected their change in VU negatively over the course of a year in Japan. The results from the simple regression analysis showed significant effects of Exp_diff on VU performance at first ($E = .07$, $t = 2.67$, $p = .01$) but not at second ($E = .02$, $t = .78$, $p = .44$) round of testing.

3.2.2. Macrostructure

The descriptive statistics of each macrostructure element is presented in Table 5. Although the total score appears to decrease from first to second round of testing, this change was not significant ($E = .12$, $t = .55$, $p = .58$). Moreover, no significant interactions between Round and Background variables (i.e. Age, L2_exp, Exp_diff, LoR) were found (p's > .44). There was also no main effect of Country ($E = .44$, $t = 1.21$, $p = .23$).

3.3. Japanese

3.3.1. Microstructure

The descriptive statistics of the Japanese microstructure are presented in Table 6. In a similar vein to the English microstructure results, there were no significant changes in Japanese microstructure over time for Type -Token Ratio (TTR) ($E = .001$, $t = .36$, $p = .71$), Typical Disfluency (TD) ($E = .16$, $t = 1.08$, $p = .28$), and Verbs per Utterance (VU) ($E = -.04$, $t = -1.39$, $p = .16$), and Mean Length of Utterance (MLU) ($E = -.33$, $t = -1.95$, $p = .06$). There was also no main effect of Country for all models (p's > .23).

Most importantly, there was a significant interaction between Round and Age at the time of return to Japan for MLU ($E = .37$, $t = 2.09$, $p = .04$) and a near-significant interaction for VU ($E = .07$, $t = 1.94$, $p = .05$). Moreover, for TTR ($E = .009$, $t = 2.02$, $p = .05$) and TD ($E = -.50$, $t = -3.15$, $p = .003$), there was a (near) significant interaction between Round and Relative Proficiency. In addition to

Figure 1. Panel (A): Interaction between Round and the Amount of reduction in L2 exposure (Exp_diff) on English Verbs per Utterance (VU) Panel (B): Interaction between Round and the Amount of reduction in L2 exposure (Exp_diff) on English Type Token Ratio (TTR).

Table 5. Descriptive statistics of English macrostructure elements for first and second round of testing.

	Action	Character	Consequence	Initiating event	Internal response	Plan	Setting	Total
Round 1	.75 (56)	2.96 (.17)	.50 (.56)	2.06 (.71)	.87 (.83)	.34 (.48)	1.71 (.92)	9.21 (2.10)
Round 2	.65 (.65)	3.0 (0)	.46 (.50)	2.06 (.84)	.71 (.72)	.25 (.43)	1.78 (1.0)	8.93 (2.21)

Relative Proficiency, Length of Residence (LoR) explained the variances in the changes of TD from first to second round of testing ($E = .47$, $t = 3.02$, $p = .005$).

The results of the linear mixed effect model show that two factors contribute to the change in Japanese microstructure between the two rounds of testing: Age and Relative Proficiency. The interactions between Age and Round for MLU and VU are presented in Figure 2 Panel (A) and (B) respectively. We see that younger children increase their MLU and VU to a greater extent than older children, but this is due to the fact that younger children start off with lower MLU and VU at the first round of testing, and thus may have had more 'room to improve' over the course of a year than older children. Indeed, simple regression analyses split by Round showed a significant effect of Age on MLU at first round ($E = 1.24$, $t = 5.25$ $p < .001$) but not at second round of testing ($E = .43$, $t = 1.56$, $p = .12$), and also a significant effect of Age on VU at first round ($E = .14$, $t = 3.02$, $p = .005$) but not at second round of testing ($E = .01$, $t = .31$, $p = .75$).

Figure 2 Panel (C) and (D) illustrate the interactions between Relative Proficiency and Round for Type-Token Ratio (TTR) and Typical Disfluency (TD) respectively. Visual inspection shows a similar trend for both TTR and TD – children who are more dominant in English at the onset of return to Japan (i.e. lower values indicate stronger dominance in English) have lower TTR and higher TD at first round of testing. However, they appear to catch up with others at the second round of testing and display the greatest improvement over time. The results from the simple regression analyses spit by Round show no significant effects of Relative Proficiency on TTR at first ($E = .01$, $t = 1.28$, $p = .20$) and second ($E = -.006$, $t = -.66$, $p = .51$) round of testing, and also no effects of Relative Proficiency on TD at first ($E = -.56$, $t = -1.69$, $p = .10$) and second ($E = .20$, $t = .59$, $p = .55$) round of testing.

In addition to the interaction between Round and Relative Proficiency, there was also a significant interaction between Round and Length of Residence (LoR) for Typical Disfluency (TD) ($E = .47$, $t = 3.02$, $p = .005$). The children who stayed longer abroad had higher TD at the first round of testing but improved their fluency to the greatest extent and performed better in the second round of testing than children who had shorter LoR abroad.

3.3.2. Macrostructure

The descriptive statistics of each macrostructure element is presented in Table 7. The total score did not increase from first to second round of testing ($E = -.45$, $t = -1.86$, $p = .07$). There was also no main

Table 6. Summary of the Japanese microstructure elements split by Round; 'Round 1' indicates children's performance at first round of testing and 'Round 2' indicates their performance at second round of testing.

	Mean	SD	Min	Max
Mean length of utterance				
Round 1	8.18	1.82	4.72	11.91
Round 2	8.80	1.58	5.28	11.65
Type-token ratio				
Round 1	.31	.04	.23	.46
Round 2	.31	.05	.21	.45
Typical disfluency				
Round 1	2.65	1.92	.22	7.32
Round 2	2.32	1.91	.20	6.45
Verbs per utterance				
Round 1	1.48	.31	.89	2.33
Round 2	1.58	.25	1.16	2.05

Figure 2. Panel (A): Interaction between Round and Age on Japanese Mean Length of Utterance(MLU) Panel (B): Interaction between Round and Age on Japanese Verbs per Utterance (VU) Panel (C): Interaction between Round and Relative Proficiency on Japanese Type Token Ratio (TTR) Panel (D): Interaction between Round and Relative Proficiency on Japanese Typical Disfluency (TD).

effect of Country ($E = .33$, $t = .56$, $p = .57$). Significant interactions between Round and background variables (i.e. Age, Exp_diff, LoR, RP) were found for Age ($E = .70$, $t = 2.78$, $p = .009$). and Relative Proficiency ($E = -.57$, $t = -2.29$, $p = .02$) only.

Table 7. Descriptive statistics of Japanese macrostructure elements for first and second round of testing.

	Action	Character	Consequence	Initiating event	Internal response	Plan	Setting	Total
Round 1	2.34 (.74)	2.37 (.70)	2.12 (.65)	2.34 (.70)	1.06 (.98)	.68 (.85)	2.68 (.73)	13.62 (3.58)
Round 2	2.65 (.54)	2.40 (.61)	2.34 (.60)	2.65 (.48)	1.03 (.89)	.50 (.84)	2.93 (.24)	14.53 (2.67)

As shown in Figure 3 Panel (A), the younger the children were at the first round of testing, the greater the improvement they show over time from first to second round of testing. Simple regression analyses split by Round reveal a non-significant effect of Age at first round of testing ($E = .43$, $t = .64$, $p = .50$) and at second round of testing ($E = -.80$, $t = -1.73$, $p = .09$). Figure 3 Panel (B) illustrates the interaction between Relative Proficiency and Round. The patterns reveal that children who were more dominant in Japanese at the onset of return to Japan show greater improvement over time, and also perform better than English-dominant children at second round of testing. Simple regression analyses split by Round reveal a non-significant effect of Relative Proficiency at first round of testing ($E = -.11$, $t = -.17$, $p = .86$) and also at second round of testing ($E = .77$, $t = 1.68$, $p = .10$).

4. Discussion

Examining the development of returnee's native heritage language as well as their L2 from the point of re-exposure to the native environment opens up new directions for research and education. First, in the context of heritage language bilingualism, while the literature shows that grammatical outcomes vary significantly across individuals and are likely to differ from monolingual and other types of native bilingual baselines (Kupisch and Rothman 2018; Montrul 2010; Polinsky 2018), findings are largely limited to contexts in which children grow up and stay in an environment where the native heritage language is a minority language. Environment transitions – as in the case of returnees – can help to tease apart various factors affecting heritage language development,

Figure 3. Panel (A): Interaction between Round and Age on Japanese total macrostructure scores Panel (B): Interaction between Round and Relative Proficiency on Japanese total macrostructure scores.

especially the roles of input and education in the heritage language. Second, the maintenance/attrition of an acquired (second) language is a topic relevant to most bilingual speakers – it is actually quite rare to find situations in which bilinguals receive sufficient amounts of 'experience' (i.e. quality and quantity of exposure, length of exposure, attitude towards the language etc.) in each language. Understanding the factors that contribute to effects of attrition (e.g. reduction in input) is not only important for revealing the nature of L2 attrition in returnees, but also provides information that can contribute to supporting the language development/maintainence of children living in a mobile international community.

Overall, we saw no significant changes in microstructure or macrostructure for both L1 and L2 over time at the group level. This finding is not too surprising, given that the body of research in both L2 attrition and heritage language reversal have shown that these processes take some time to surface at the behavioral level. For instance, Yoshitomi's (1999) longitudinal study on four returnee children over the course of a year showed very little signs of attrition in the domains of lexicon (TTR), fluency, verb morphology, and articles. Also, in Tomiyama's (1999, 2000) study, fluency is the only aspect that is vulnerable to effects of attrition in the first 19 months of the child's stay in Japan. Alternatively, the re-exposure effect of the heritage language seems to only appear after a prolonged period of stay in the heritage language dominant environment of seven to eight years (Daller, Treffers-Daller, and Furman 2011; Daller and Yıldız 1995; Treffers-Daller, Özsoy, and Van Hout 2007, 2016). However, we should note that proportion of increase in Japanese Mean Length of Utterance (MLU) ($p = .06$) and total macrostructure scores ($p = .07$) were approaching significance, while the changes observed in English were far from reaching significance for both microstructure (p's > .44) and macrostructure ($p > .58$). This trend suggests that returnee children as a group were able to maintain their L2 English narrative abilities while (nearly) improving their Japanese narrative abilities over time.

4.1. English microstructure and macrostructure

Interestingly, the present results show that certain bilingual factors explain individual variability observed from first to second round of testing. Indeed, we used these factors to understand not only individual difference *par excellence*, but individual difference in a longitudinal sense from each participant's own baseline. These are important steps in the right direction given that, ultimately, what we seek to capture and explain in these understudied bilingual populations is not only what they do as cohorts, but the extent to which they vary from one another and, if applicable, the systematicity underlying individual differences (Kupisch and Rothman 2018). By engaging with variables that proxy for various types of bilingual experiences and having a high degree of ecological validity via longitudinal baselines, the present study accords nicely with recent trends in bilingualism research calling for a 'bilingual turn' in control measures and acknowledging the dynamic nature and complexities of engagement with bilingual experiences that are likely to have individual-level knock-on effects (e.g. DeLuca, Rothman, and Pliatsikas 2019; Gullifer and Titone 2019; Luk and Bialystok 2013; Ortega 2009).

An individual'a rate of decline in L2 exposure predicted changes in Type-Token Ration (TTR) and Verbs per Utterance (VU) over time. The patterns of interactions in Figure 1 indicate that the individual children who experienced the most decline in their L2 exposure showed the greatest decrease in their English TTR and VU over the course of a year. It is interesting to note that the cross-over point between children who exhibited an increase, as opposed to a decrease in their performance for TTR and VU, both lie around the 40% to 45% mark of loss in L2 exposure. That is, children who experienced at least 40% of reduction in L2 exposure show effects of attrition in their TTR and VU. This finding is in fact in line with a study by (Kubota, Chevalier, and Sorace 2020b) which demonstrated that losing 40% or more exposure to the L2 may have attenuated effects of bilingualism on their development in executive function.

The present data provide additional credence to (at least) 40% reduction in (L2) exposure as being a critical mass tipping point for maintenance of bilingual effects, be them linguistic or cognitive in

nature. Of course, much more research is warranted to determine if this (at least) 40% threshold is at all generalizable. Given the high theoretical and practical significance of revealing what the threshold of minimal exposure for maintenance is, the fact that the divide we see in our data is not random and, better yet, overlaps with the same percentage found in Kubota, Chevalier, and Sorace (2020b), is promising. On the theoretical side, knowing what levels of input, among other factors, are required for linguistic maintenance can help adjudicate between competing models of attrition, on the one hand, as well as shed some (unique) light on the underlying mechanisms. From a practical standpoint, knowing this same information helps to build bridges between research and its stakeholders, informing recommendations and best practices for returnees, their parents, educators and policy makers who wish to know what variables within their control they can manipulate for a desired outcome.

It is somewhat unexpected that L2 exposure affected the rate of attrition in lexical diversity (TTR) and syntactic complexity (VU), since these two aspects of language have been previously found to gradually decrease over time in qualitative case-studies (Tomiyama 1999, 2000, 2008; Yoshitomi 1999). Changes in fluency and MLU, on the other hand, were not predicted by any of the bilingual variables, although fluency is expected to be one of the most vulnerable linguistic aspects for attrition during the initial stages of return to the native language environment (Hansen, Gardner, and Pollard 1998; Tomiyama 1999). Fluency, however, is a variable that can be measured in several distinct ways, as opposed to MLU or TTR in which its calculations are more consistent across studies. For instance, fluency can be measured by unfilled pause time and frequency (Hansen, Gardner, and Pollard 1998), pauses, repetitions, and self-repairs (Tomiyama 1999, 2000), or words per minute (Yoshitomi 1999). Such variability in the measurement of fluency may have contributed to the discrepancy between our study and previous qualitative work.

In contrast to the effects of continued L2 exposure on changes in English microstructure over time, no bilingual variables explained the variance observed in the changes in English macrostructure. Moreover, contrary to our predictions, English macrostructure did not improve over time, but rather remained relatively stable. In the context of L2 attrition, exposure may have a greater impact on the changes in microstructure than macrostructure, given that microstructure – especially lexically dependent measures – has been found to be more sensitive to effects of input and exposure in the bilingual narrative literature (Gagarina 2012; Iluz-Cohen and Walters 2012; Rodina 2017).

4.2. Japanese microstructure and macrostructure

In terms of the returnee children's narrative performance in Japanese, age at return/age at first round of testing predicted the development of both micro- and macrostructural elements. Specifically, the younger the children were at the point of return to Japan, the greater they improved in their Japanese knowledge of grammar, namely Mean Length of Utterance (MLU) and Verbs per Utterance (VU). However, it is important to note here that as shown in Figure 2 (Panel (A) and (B)), the interaction between Round and Age is mainly motivated by the slope in the first round of testing, indicating that younger children, unsurprisingly, had significantly lower MLU and VU than older children at the point of return to Japan. Therefore, the younger children had more room to mature and improve their performance over time than older children. Interestingly, the threshold age for exhibiting positive changes in MLU, VU, and macrostructure was around 10.8 years old. Children whose ages were below this threshold exhibited significant gains in their L1 grammar as well as macrostructure. In particular, children around the ages of 7;5 years (90 months) to 8;3 years (100 months) experienced the greatest increase in narrative abilities over time. Our results support the findings of other longitudinal studies that showed a significant increase in micro- and macrostructure among kindergarten to first-grade Spanish-English bilingual children (Bitetti and Hammer 2016; Squires et al. 2014; Uccelli and Páez 2007). A longitudinal case study over 14 years by Taura and Taura (2012) also found that Japanese-English bilingual's story grammar showed significant gains from ages 4;09 to 7;00 and plateaued around 11;0 (although this was in their L2 English). The rapid gains in micro- and

macrostructure around the ages of 6 to 8 most likely coincides with the introduction of formal schooling and literacy practices (Bitetti and Hammer 2016).

Another experiential factor that contributed to explaining returnee children's Japanese narrative abilities over time was relative proficiency, measured by utilizing the verbal fluency task. In terms of their microstructure, relative proficiency affected development in Typical Disfluency (TD) and Type-Token Ratio (TTR). The more dominant they were in English at the onset of return to Japan, the better they developed their Japanese fluency and lexical diversity. It appears to be the case that children who were more dominant in English (and thus less dominant in Japanese) benefited the most from the environmental shift in language environment. As shown in Figure 2 (Panel (C) and (D)), although no significant simple effects of relative proficiency were found for each round of testing, it is clear that the slope of first round of testing is steeper than second round of testing for both TTR and TD. This means that the returnee children who were more proficient in English had lower TTR and higher TD in their L1 Japanese upon returning to Japan. However, after a year of immersion in Japan, they seem to catch up to their peers who were already dominant in Japanese or balanced in their two languages. Additionally, children who lived longer in the L2 dominant environment better improved their fluency over time as indicated by the interaction between Round and Length of Residence (LoR). Indeed, LoR has been found to be associated with narrative abilities in bilingual children (Govindarajan and Paradis 2019; Hao et al. 2019). Taken together, the findings indicate that children who had less proficiency and limited length of exposure in Japanese seem to take advantage of the re-exposure effects in the L1/HL and eventually perform comparably to their counterparts after a year of re-immersion in the native language environment. This suggests that it is crucial for children with lower proficiency and limited experience in their heritage language/native language, which could potentially be a majority of returnees, to receive language-related support especially in their first year of returning, since this appears to be an important time-window in which exponential development in their L1 takes place.

While children who were more proficient in English improved their Japanese microstructure (TTR and TD) over time, the relationship between relative proficiency and macrostructure suggests the opposite – that is, children who were more proficient in Japanese improved their Japanese macrostructure to a greater extent. Since macrostructure refers to global and higher-order conceptual (and linguistic) knowledge and thus may be less sensitive to input effects than microstructure, perhaps a certain level of Japanese proficiency is needed for macrostructure to benefit from increased L1 input.

To summarize, our longitudinal study investigated the changes in returnee children's L1 Japanese and L2 English in a unique context in which the reversal of language status occurs from the principle language of exposure to the minority one, inversely for English and Japanese. Although no significant group-level changes were found in either L1 or L2 upon return to Japan, we found that various experiential factors related to bilingualism proved explanatory for variances in microstructure and macrostructure over time. Our study is the first to track both languages of bilingual returnee children over time over the course of a year, capturing their language development immediately after their return to Japan. Given that effects of L2 attrition and heritage language reversal appear to take some time to manifest in the returnee's grammar, future research that tracks both languages of returnee children for a longer period of time will inform us further about their developmental trajectories and what factors influence these processes.

Note

1. It is helpful to contextualize the specific reality of our returnee families outside of an English-majority language society. These returnee families are of a transient type, that is, they know from the time they leave Japan not only that they will return but exactly for how long they will be abroad (related to long periods of secondment for work). And so, as a matter of inclusion these families made a conscious decision to enroll their children in

English-only international schools. As a result, the social networks of these returnee children are primarily made up of either other Japanese immigrant families or other international families whereby Japanese and English are the overwhelmingly available lingua francas, as we will see this bears out in our data (i.e. lack of difference between them and those having lived in the English-majority settings).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by British Association for Japanese Studies [grant number JC2018-6-1]; PPLS Small Research Project Fund; Great Britain Sasakawa Foundation [grant number 5038]; Tromsø forskningsstiftelse; Japan Foundation Endowment Committee [grant number 5780416].

Notes on contributors

Maki Kubota is a Postdoctoral Researcher in the HeLPiNG (Heritage-bilingual Linguistic Proficiency in their Native Grammar) project at the Department of Language and Culture at UiT, and the lab's Outreach Coordinator. Her work mainly involves examining language development in children, including bilinguals, attriters (returnees), and heritage language speakers.

Vicky Chondrogianni is a Senior Lecturer in the School of Philosophy, Psychology and Language Sciences at the University of Edinburgh. Her research focuses on cross-linguistic aspects of acquisition and processing in typically-developing monolingual and bilingual children and in children with language impairment.

Adam Scott Clark is a curriculum developer based in Toronto, Canada. His research focuses on bilingual education and policy, particularly in regions where there are two or more official languages.

Jason Rothman is Professor of Linguistics at UiT and Adjunct Professor at Universidad Nebrija (Spain). He primarily works on language acquisition and processing across the life span, especially in various types of bilingualism and multilingualism.

ORCID

Maki Kubota http://orcid.org/0000-0003-0580-1045

References

Aktas, E., and S. U. Yurt. 2017. "Effects of Digital Story on Academic Achievement, Learning Motivation and Retention among University Students." *International Journal of Higher Education* 6 (1): 180–196.

Altman, C., S. Armon-Lotem, S. Fichman, and J. Walters. 2016. "Macrostructure, Microstructure, and Mental State Terms in the Narratives of English-Hebrew Bilingual Preschool Children with and Without Specific Language Impairment." *Applied Psycholinguistics* 37 (1): 165–193.

Bailey, A. L., A. C. Moughamian, K. R. Kelly, A. McCabe, and B. H. Huang. 2018. "Leap-frog to Literacy: Maternal Narrative Supports Differentially Relate to Child Oral Language and Later Reading Outcomes." *Early Child Development and Care*, 190(7), 1136-1149.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. "Fitting linear mixed-effects models using lme4." *Journal of Statistical Software* 67 (1): 1–48.

Bitetti, D., and C. S. Hammer. 2016. "The Home Literacy Environment and the English Narrative Development of Spanish–English Bilingual Children." *Journal of Speech, Language, and Hearing Research* 59 (5): 1159–1171.

Daller, M. H., J. Treffers-Daller, and R. Furman. 2011. "Transfer of Conceptualization Patterns in Bilinguals: The Construal of Motion Events in Turkish and German." *Bilingualism: Language and Cognition* 14 (1): 95–119.

Daller, H., and C. Yıldız. 1995. "Die türkischen Sprachfähigkeiten von Rückkehrern aus Deutschland." *Zwischen Den Sprachen: Sprachgebrauch, Sprachmischung Und Sprachfähigkeiten Türkischer Rückkehrer Aus Deutschland* 2: 83–94.

DeLuca, V., J. Rothman, and C. Pliatsikas. 2019. "Linguistic Immersion and Structural Effects on the Bilingual Brain: A Longitudinal Study." *Bilingualism: Language and Cognition* 22 (5): 1160–1175.

Eurostat. 2019. *Migration and migrant population statistics*. https://ec.europa.eu/eurostat/statistics-explained/pdfscache/1275.pdf.

Flores, C. 2010. "The Effect of age on Language Attrition: Evidence from Bilingual Returnees." *Bilingualism: Language and Cognition* 13 (4): 533–546.
Flores, C. 2019. "Language Development of Bilingual Returnees." In *Oxford Handbook of Language Attrition*, edited by M. S. Schmid and B. Köpke, 493–501. Oxford: Oxford University Press.
Flores, C. 2020. "Attrition and Reactivation of a Childhood Language: The Case of Returnee Heritage Speakers." *Language Learning* 70 (1): 85–121.
Flores, C., and A. Rato. 2016. "Global Accent in the Portuguese Speech of Heritage Returnees." *Heritage Language Journal* 13 (2): 161–183.
Flores, C., and N. Snape. Forthcoming. "Language Attrition and Heritage Language Reversal in Returnees." In *The Cambridge Handbook of Heritage Languages and Linguistics*, edited by S. Montrul and M. Polinsky. Cambridge University Press.
Gagarina, N. 2012. "Discourse Cohesion in the Elicited Narratives of Early Russian-German Sequential Bilinguals." In *Multilingual Individuals and Multilingual Societies*, edited by K. Braunmüller and C. Gabriel, 101–120. Amsterdam: John Benjamins.
Gagarina, N., D. Klop, I. M. Tsimpli, and J. Walters. 2016. "Narrative Abilities in Bilingual Children." *Applied Psycholinguistics* 37 (1): 11–17.
Gillam, S., R. Gillam, J. Fargo, A. Olszewski, and H. Segura. 2017. "Monitoring Indicators of Scholarly Language (MISL): A Progress-Monitoring Tool for Scoring Narratives Produced by School-age Children with Language Impairments." *Communication Disorders Quarterly* 38 (2): 96–106.
Govindarajan, K., and J. Paradis. 2019. "Narrative Abilities of Bilingual Children with and Without Developmental Language Disorder (SLI): Differentiation and the Role of Age and Input Factors." *Journal of Communication Disorders* 77: 1–16.
Gullifer, J. W., and D. Titone. 2019. "Characterizing the Social Diversity of Bilingualism Using Language Entropy." *Bilingualism: Language and Cognition* 94 (1): 1–18.
Hansen-Strain, L. 1990. "The Attrition of Japanese by English-Speaking Children: An Interim Report." *Language Sciences* 12 (4): 367–377.
Hansen, L., J. Gardner, and J. Pollard. 1998. "The Measurement of Fluency in a Second Language: Evidence from the Acquisition and Attrition of Japanese." *On JALT* 97: 37–45.
Hao, Y., L. M. Bedore, L. Sheng, and E. D. Peña. 2019. "Narrative Skills in two Languages of Mandarin–English Bilingual Children." *International Journal of Speech-Language Pathology* 21 (4): 325–335.
Iluz-Cohen, P., and J. Walters. 2012. "Telling Stories in Two Languages: Narratives of Bilingual Preschool Children with Typical and Impaired." *Bilingualism: Language and Cognition* 15 (1): 58–74.
Justice, L. M., R. P. Bowles, J. N. Kaderavek, T. A. Ukrainetz, S. L. Eisenberg, and R. B. Gillam. 2006. "The Index of Narrative Microstructure: A Clinical Tool for Analyzing School-age Children's Narrative Performances." *American Journal of Speech-Language Pathology* 15 (2): 177–191.
Kapalková, S., K. Polisenska, L. Marková, and J. Fenton. 2016. "Narrative Abilities in Early Successive Bilingual Slovak-English Children: A Cross-Language Comparison." *Applied Psycholinguistics* 37 (1): 145–164.
Kubota, R. 2012. "Language and Education for Returnees." In *The Encyclopedia of Applied Linguistics*, edited by C. A. Chapelle. doi:10.1002/9781405198431.
Kubota, M., N. Chevalier, and A. Sorace. 2020a. "How Bilingual Experience and Executive Control Influence Development in Language Control among Bilingual Children." *Developmental Science* 23 (1): e12865.
Kubota, M., N. Chevalier, and A. Sorace. 2020b. "Losing Access to the Second Language and its Effect on Executive Function Development in Childhood: The Case Of'returnees." *Journal of Neurolinguistics* 55: 100906.
Kunnari, S., T. Välimaa, and P. Laukkanen-Nevala. 2016. "Macrostructure in the Narratives of Monolingual Finnish and Bilingual Finnish-Swedish Children." *Applied Psycholinguistics* 37 (1): 123.
Kupisch, T., and J. Rothman. 2018. "Terminology Matters! Why Difference is Not Incompleteness and How Early Child Bilinguals are Heritage Speakers." *International Journal of Bilingualism* 22 (5): 564–582.
Loban, W. 1976. *Language Development: Kindergarten Through Grade Twelve. Research Report No. 18*. Urbana, IL: National Council of Teachers of English.
Lucero, A. 2015. "Cross-linguistic Lexical, Grammatical, and Discourse Performance on Oral Narrative Retells among Young Spanish Speakers." *Child Development* 86 (5): 1419–1433.
Luk, G., and E. Bialystok. 2013. "Bilingualism is not a Categorical Variable: Interaction between Language Proficiency and Usage." *Journal of Cognitive Psychology* 25 (5): 605–621.
MacWhinney, B. 2000. *The CHILDES Project: Tools for Analyzing Talk*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.
Mayer, M. 1969. *Frog, Where are you?* New York: Dial Press.
Mayer, M. 1973. *Frog on his Own*. New York: Dial Press.
Méndez, L. I., J. Perry, Y. Holt, H. Bian, and S. Fafulas. 2018. "Same or Different: Narrative Retells in Bilingual Latino Kindergarten Children." *Bilingual Research Journal* 41 (2): 150–166.
Miller, J. F., J. Heilmann, A. Nockerts, A. Iglesias, L. Fabiano, and D. J. Francis. 2006. "Oral Language and Reading in Bilingual Children." *Learning Disabilities Research & Practice* 21 (1): 30–43.
Montrul, S. 2010. "Current Issues in Heritage Language Acquisition." *Annual Review of Applied Linguistics* 30: 3–23.

O'neill, D. K., M. J. Pearce, and J. L. Pick. 2004. "Preschool Children's Narratives and Performance on the Peabody Individualized Achievement Test–Revised: Evidence of a Relation between Early Narrative and Later Mathematical Ability." *First Language* 24 (2): 149–183.
Ortega, L. 2009. *Understanding Second Language Acquisition*. London: Hodder Education.
Pesco, D., and E. K.-R. Bird. 2016. "Perspectives on Bilingual Children's Narratives Elicited with the Multilingual Assessment Instrument for Narratives." *Applied Psycholinguistics* 37 (1): 1–9.
Polinsky, M. 2018. *Heritage Languages and Their Speakers*. Cambridge: Cambridge University Press.
Reetz-Kurashige, A. 1999. "Japanese Returnees' Retention of English-Speaking Skills: Changes in Verb Usage Over Time." In *Second Language Attrition in Japanese Contexts*, edited by L. Hansen, 21–58. Oxford: Oxford University Press.
Roch, M., E. Florit, and C. Levorato. 2016. "Narrative Competence of Italian-English Bilingual Children between 5 and 7 Years." *Applied Psycholinguistics* 37 (1): 49.
Rodina, Y. 2017. "Narrative Abilities of Preschool Bilingual Norwegian-Russian Children." *International Journal of Bilingualism* 21 (5): 617–635.
Rothman, J. 2009. "Understanding the Nature and Outcomes of Early Bilingualism: Romance Languages as Heritage Languages." *International Journal of Bilingualism* 13 (2): 155–163.
Sekiguchi, T. 2003. *Zainichi nikkeijin wa ima: "Kyoukai kuukan ni ikiru kodomotachi" no shiten kara, nihon no gakkou o kouryo suru: Zainichi nikkei burajirujin no jirei o chuushin to shite—Part II*. Tokyo: Kikan Kaigai Nikkeijin.
Simon-Cereijido, G., and V. F. Gutiérrez-Clellen. 2009. "A Cross-Linguistic and Bilingual Evaluation of the Interdependence between Lexical and Grammatical Domains." *Applied Psycholinguistics* 30 (2): 315.
Singmann, H., B. Bolker, J. Westfall, and F. Aust. 2015. afex: Analysis of factorial experiments. *R Package Version 0.13–145*.
Spencer, T. D., and D. B. Petersen. 2018. "Bridging Oral and Written Language: An Oral Narrative Language Intervention Study with Writing Outcomes." *Language, Speech, and Hearing Services in Schools* 49 (3): 569–581.
Squires, K. E., M. J. Lugo-Neris, E. D. Peña, L. M. Bedore, T. M. Bohman, and R. B. Gillam. 2014. "Story Retelling by Bilingual Children with Language Impairments and Typically Developing Controls." *International Journal of Language & Communication Disorders* 49 (1): 60–74.
Taura, H. 2001. "L2 Attrition&Maintenance Observed in Japanese Returnee Students in Terms of Productive Oral Proficiency." *Journal of Fukui Medical University* 2 (1): 71–79.
Taura, H., and A. Taura. 2000. "Reverse Language Attrition Observed in Japanese Returnee Students' English Productive Skills." *Journal of Fukui Medical University* 1 (3): 535–543.
Taura, H., and A. Taura. 2012. "Linguistic and Narrative Development in a Japanese–English Bilingual's First Language Acquisition: A 14-Year Longitudinal Case Study." *International Journal of Bilingual Education and Bilingualism* 15 (4): 475–508.
To, C. K.-S., S. F. Stokes, H.-T. Cheung, and B. T'sou. 2010. "Narrative Assessment for Cantonese-Speaking Children." *Journal of Speech, Language, and Hearing Research* 53 (3): 648–669.
Tomiyama, M. 1999. "The First Stage of Second Language Attrition: A Case Study of a Japanese Returnee." In *Second Language Attrition in Japanese Contexts*, edited by L. Hansen, 59–79. Oxford: Oxford University Press.
Tomiyama, M. 2000. "Child Second Language Attrition: A Longitudinal Case Study." *Applied Linguistics* 21 (3): 304–332.
Tomiyama, M. 2008. "Age and Proficiency in L2 Attrition: Data from two Siblings." *Applied Linguistics* 30 (2): 253–275.
Treffers-Daller, J., M. Daller, R. Furman, and J. Rothman. 2016. "Ultimate Attainment in the use of Collocations among Heritage Speakers of Turkish in Germany and Turkish–German Returnees." *Bilingualism: Language and Cognition* 19 (3): 504–519.
Treffers-Daller, J., A. S. Özsoy, and R. Van Hout. 2007. "(In) Complete Acquisition of Turkish among Turkish–German Bilinguals in Germany and Turkey: An Analysis of Complex Embeddings in Narratives." *International Journal of Bilingual Education and Bilingualism* 10 (3): 248–276.
Uccelli, P., and M. M. Páez. 2007. "Narrative and Vocabulary Development of Bilingual Children from Kindergarten to First Grade: Developmental Changes and Associations among English and Spanish Skills." *Language, Speech, and Hearing Services in Schools* 38 (3): 225–236.
Unsworth, S. 2016. "Amount of Exposure as a Proxy for Dominance in Bilingual Language Acquisition." In *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*, edited by C. Silva-Corvalán and J. Treffers-Daller, 156–173. Cambridge: Cambridge University Press.
Westerveld, M. F. 2014. "Emergent Literacy Performance Across Two Languages: Assessing Four-Year-old Bilingual Children." *International Journal of Bilingual Education and Bilingualism* 17 (5): 526–543.
Yang, Y.-T. C., and W.-C. I. Wu. 2012. "Digital Storytelling for Enhancing Student Academic Achievement, Critical Thinking, and Learning Motivation: A Year-Long Experimental Study." *Computers & Education* 59 (2): 339–352.
Yoshida, T., D. Matsumoto, S. Akashi, T. Akiyama, A. Furuiye, C. Ishii, and N. Moriyoshi. 2009. "Contrasting Experiences in Japanese Returnee Adjustment: Those who Adjust Easily and Those who do not." *International Journal of Intercultural Relations* 33 (4): 265–276.
Yoshitomi, A. 1999. "On the Loss of English as a Second Language by Japanese Returnee Children." In *Second Language Attrition in Japanese Contexts*, edited by L. Hansen, 80–111. Oxford University Press.